Article

A Sparse Manifold Classification Method Based on a Multi-Dimensional Descriptive Primitive of Polarimetric SAR Image Time Series

Chu He [1,2,*], Gong Han [1], Di Feng [1], Juan Du [3,*] and Mingsheng Liao [2]

[1] Signal Processing Laboratory, Electronic Information School, Wuhan University, Wuhan 430072, China; hangong@whu.edu.cn (G.H.); fengdi@whu.edu.cn (D.F.)
[2] State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; liao@whu.edu.cn
[3] Remote Sensing and Information Engineering School, Wuhan University, Wuhan 430079, China
[*] Correspondence: chuhe@whu.edu.cn (C.H.); dujuan_rs@whu.edu.cn (J.D.); Tel.: +86-133-0716-2028 (C.H.); +86-186-0272-0952 (J.D.)

Abstract: Classification using the rich information provided by time-series and polarimetric Synthetic Aperture Radar (SAR) images has attracted much attention. The key point is to effectively reveal the correlation between different dimensions of information and form a joint feature. In this paper, a multi-dimensional SAR descriptive primitive for each single pixel is firstly constructed, which in the polarimetric scale obtains incoherent information through target decompositions while in the time scale obtains coherent information through stochastic walk. Secondly, for the purpose of feature extraction and dimension reduction, a special feature space mapping for the descriptive primitive of the whole image is proposed based on sparse manifold expression and compressed sensing. Finally, the above feature is inputted into a support vector machine (SVM) classifier. This proposed method can inherently integrate the features of polarimetric SAR times series. Experiment results on three real time-series polarimetric SAR data sets show the effectiveness of our presented approach. The idea of a multi-dimensional descriptive primitive as a convenient tool also opens a new spectrum of potential for further processing of polarimetric SAR image time series.

Keywords: polarimetric SAR time series; image classification; multi-dimensional descriptive primitive; sparse manifold expression; compressed sensing

1. Introduction

Synthetic aperture radar (SAR) can be used in a vast majority of application areas due to its ability to work day and night under all weather conditions. However, its coherent imaging mechanism and extremely low signal-to-noise ratio make it difficult to interpret. In recent years, increasing access of polarimetric and time-series SAR data has provided abundant information regarding the specific location. However, how to fully utilize this information has emerged as a basic problem in SAR image interpretation.

With respect to the incoherent information from a single SAR image, quite a few product statistical model distributions have been proposed for classification. Classical ones include the earliest Gamma distribution, K distribution, G0 distribution and G distribution [1]. Note that the use of polarimetry for radar remote sensing is increasingly extensive; a set of methods known collectively as target decomposition (TD) theorems springs up, which were first formalized by Huynen [2] but have their roots in the work of Chandrasekhar on light scattering by small anisotropic particles [3]. Since this

original work, there have been many other effective decompositions such as Cloude decomposition, Holm decomposition, Krogager decomposition and Huynen decomposition [4].

As for the coherent information from multiple images, time-series SAR provides the possibility of extracting the interference information. By unwrapping the achieved information, the elevation information and structure information of the ground can be obtained. Representative instances are Branch-Cut Algorithm, Minimum Discontinuity, Mask-Cut Algorithm and Minimum L^p-Norm Phase Unwrapping [5].

With this incoherent and coherent information, researchers have developed various image processing patterns for SAR classification from the combination of polarimetric distribution and classifiers [6] to the combination of extracted features and classifiers [7] and nowadays to deep learning [8]. Applicable to all patterns, a promising direction is to optimally relate the incoherent information together with the coherent information.

So far, many feature-fusing methods have been proposed. Here, we introduce three kinds of processing sketches: co-training, multiple kernel learning and subspace learning [9–12]. During each iteration of Co-training [13], models are trained separately but relative to each feature and then the algorithm propagates the disagreement of the two models back to the training set. In multiple kernel learning methods, each feature corresponds to a kernel which best matches its property and then these kernels are simultaneously connected in a linear or nonlinear way. One representative multiple kernel learning algorithm is Simple MKL which combines kernels linearly and sparsely [14]. Subspace learning aims at finding the shared underlying subspace on the assumption that the original features are generated from this latent subspace by a specific mapping. Principle Components Analysis (PCA) [15] is a time-honored and simple technique performing subspace learning. Figure 1 sketches the three processing approaches.
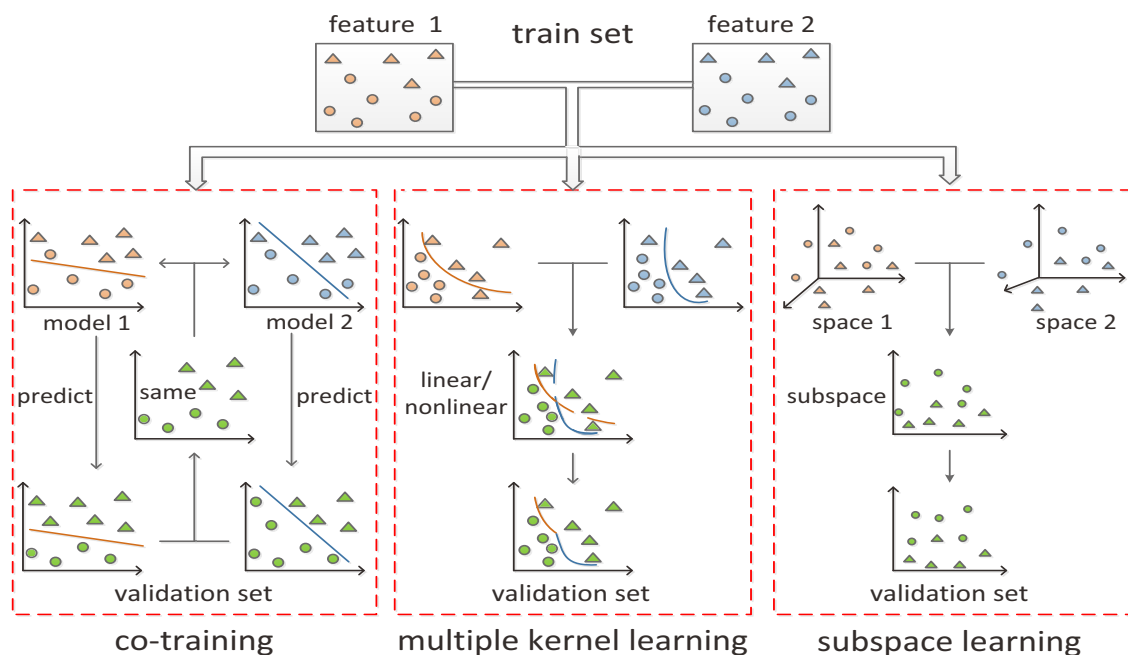
Figure 1. Sketches of co-training, multiple kernel learning and subspace learning.

However, the involvement of unwrapping and the complexity of the InSAR inverse procedure complicate the process of accurate height inversion for classification, making how to obtain coherent information without unwrapping and inversion a critical issue. Worse still, neither co-training, nor multiple kernel learning nor subspace learning merge features at the initial feature construction stage. In this framework, we are dedicated to exploring a rather concise approach which can also inherently integrate the features of polarimetric SAR times series.

The main contributions of our work are two-fold. Firstly a multi-dimensional SAR descriptive primitive for each single pixel is constructed, which in the polarimetric scale obtains incoherent information, while in the time scale obtains coherent information. The descriptive primitive can become a convenient tool for the processing of polarimetric SAR image time series. Secondly, considering the inconsistency between polarimetric incoherent scale and time-series coherent scale, a nonlinear classification model is further constructed based on sparse manifold expression and compressed sensing for feature extraction and dimension reduction. This model can deal with the inconsistency of the two scales and tactfully avoid the nonlinearity problem brought by the multiplicative model of SAR.

The rest of the paper is organized as follows. Section 2 is dedicated to the creation of a multi-dimensional descriptive primitive. Section 3 introduces the sparse manifold classification model. Section 4 shows the validations on three real polarimetric SAR data sets. The conclusion is included in Section 5.

2. The Multi-Dimensional Descriptive Primitive

2.1. Incoherent Feature in the Polarization Scale

In the polarimetric scale, target decompositions are taken advantages of to generate incoherent information of every single SAR image, which include Pauli decomposition, SDH decomposition, Huynen decomposition, Holm decomposition and Cloude decomposition in this paper. Every decomposition creates three parameters for a single pixel. We bunch together all these parameters generated by these five decompositions into a 5×3 matrix.

2.2. Coherent Feature in the Time Scale

In the time scale, stochastic walk is utilized to form the coherent feature of the time-series data. Stochastic walk was first used by Francisco Estrada et al. for image denoising [16]. Its simple concept is set on the basis of random walk probabilities. Every random walk, beginning from a given pixel, paths smoothly over arbitrary surrounding neighborhoods. The random walk probability between paired pixels is determined by their similarity, which serves as a weight.

Define an ordered pixel sequence $T_{0,k} = \{x_0, x_1, \ldots, x_k\}$ to represent a path from x_0 to x_k. The transition probability between two consecutive pixels x_j and x_{j+1} within this sequence is inversely proportional to both the dissimilarity between x_j and x_{j+1} and the dissimilarity between x_0 and x_{j+1} [16], which is expressed by Equation (1).

$$p\left(x_{j+1} \mid x_j\right) = \frac{1}{K} e^{\frac{-d\left(x_0, x_{j+1}\right)^2}{2\delta^2}} e^{\frac{-d\left(x_j, x_{j+1}\right)^2}{2\delta^2}} \tag{1}$$

Here, K is a parameter for normalization and δ is for scaling. $d\left(x_i, x_j\right)$ is a dissimilarity measure relating image pixel x_i and x_j. According to the first-order Markov assumption, the probability of the whole sequence is accessible. With $T_{0,k}$ in hand, $T_{0,k+1}$ can be directly created by generating the neighborhood of x_k and then selecting a neighbor with probability $p\left(x_{k+1} \mid x_k\right)$. If m random walks are given and each is originating from x_0 with a length of k steps, the final result shall be calculated as the weighted mean of all pixels visited in every step during every walk.

In the proposed method, the original 2-D neighborhood is expanded into a 3-D one. The neighborhood of a specific pixel in time t consists of not only the eight pixels surrounding it but also the 18 pixels in adjacent time $t-1$ and $t+1$.

2.3. Multi-Dimensional Descriptive Primitive

In registered polarimetric SAR images, each single pixel has its corresponding polarimetric matrix T. Firstly, extract all the 3×3 matrix T of a same point on different dates and connect all

the matrixes together according to the time sequence. This is a prototype of our multi-dimensional descriptive primitive. On this primitive, we implement target decompositions in the polarimetric scale and then stochastic walk in the time scale, thus integrating the incoherent feature with the coherent feature to get our multi-dimensional descriptive primitive of a single point. In our work, we selected three discrete dates to generate a 3-D descriptive primitive which is shown in Figure 2. The three 3×3 matrix T of the same point on different dates, together, made up a three-dimensional cube, which was the raw material of the subsequent processes. In the polarimetric scale, each 3×3 matrix T was dealt with 5 target decompositions. With every decomposition generating 3 parameters, three 3×3 matrix T were transformed into three 5×3 incoherent feature matrixes, constructing a primitive with only incoherent feature. In the time scale, stochastic walk was utilized on the three 5×3 matrixes to update the incoherent feature primitive with the coherent feature. Then, all the bases were connected together to generate a three-dimensional descriptive primitive of a whole graph.
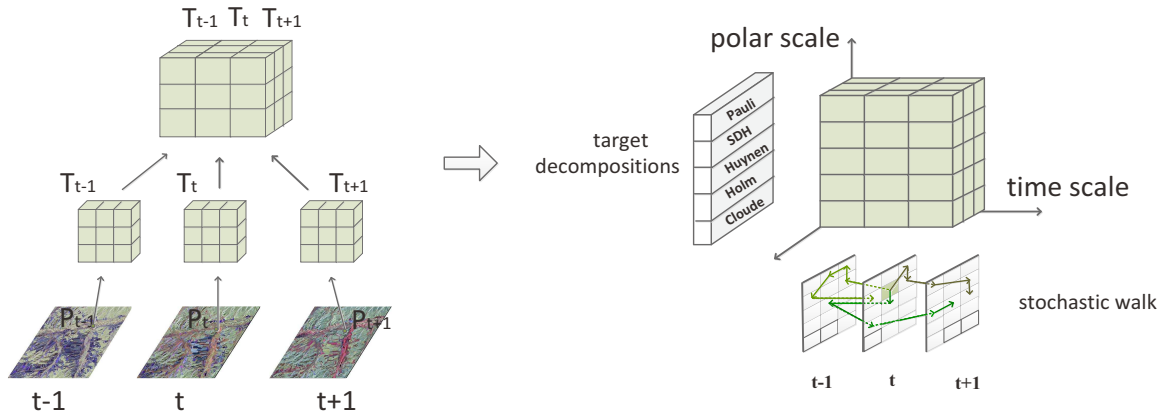
Figure 2. A 3-D descriptive primitive of a single point.

3. The Sparse Manifold Classification Model

When the size of the involved SAR image is relatively big, the problem may occur that the descriptive primitive is too large for the subsequent operations. Besides, the problem of the nonlinearity multiplicative model brought by the coherent imaging mechanism of SAR calls for an adaptable approach. To these ends, we propose a nonlinear classification model based on sparse manifold expression and compressed sensing.

3.1. Sparse Manifold Expression

Sparse representation has always been an effective feature extraction approach in the classification area. Existing sparse coding models linearly assemble the basic atoms in an over-complete dictionary to approximate the input signal. Assuming that x_i is the d dimensional local feature of the SAR image, i.e., $X = [x_1, \ldots, x_N] \in R^{d \times N}$ and $B = [b_1, \ldots, b_M] \in R^{d \times M}$ is the dictionary with M entries, every input can be represented by its most similar M-dimensional code ω_i which satisfies $x_i = B\omega_i$. $\Omega = [\omega_1, \ldots, \omega_N]$ is the set of the codes. Taking the Locality-constrained Linear Coding (LLC) model [17] as an example, we can break it into two parts, the first of which is a coding error term and captures typical features of local description; the second part constraints the code to make sure that the under-determined system of the equation has a unique solution. The following expression gives the details.

$$\underset{\Omega}{\arg\min} \sum_{i=1}^{N} \|x_i - B\omega_i\|^2 + \lambda \|d_i \cdot \omega_i\|^2 \quad s.t. \; 1^T w_i = 1, \; \forall i \tag{2}$$

λ is a constraint parameter and "$\cdot$" denotes the multiplication in element-wise. $d_i \in R^M$ represents the locality adaptor to maintain the similarity between the base vector and the input

descriptor, which is usually normalized to be between $(0, 1]$. The constraint $1^T w_i = 1$ guarantees the shift-invariant requirements of the LLC code. Though the LLC model performs well in most situations, it is not applicable to nonlinear cases. In our context, to better integrate the polarimetric scale and the time-series scale and to handle the nonlinearity of the SAR multiplicative model, we bring in sparse manifold expression to find the potential low dimensional structure in the feature space. Manifold learning supposes that points in a high dimensional space virtually exist in a low dimensional manifold. Inspired by Locally Linear Embedding (LLE) [18], we resort to a manifold by preserving the neighborhood in the original data space and then mapping the data into global internal coordinates on this manifold while keeping the local geometry unchanged. The data descriptor with the form of a 3-D primitive in this paper is no longer $x \in R^d$ but $x' \in R^{5 \times 3 \times t}$. t is the number of different dates in the time-series data set. In our experiments, each data set has three different dates, i.e., $t = 3$. We reconstruct each x' from its neighbors by linear coefficients $\gamma_{i,j}$ in Equation (3) and then fix the coefficients to optimize y_i in the low-dimensional manifold which corresponds to x' in Equation (4).

$$\underset{\Gamma}{\text{argmin}} \sum_i \|x_i' - \sum_j \gamma_{i,j} x_j'\|^2 \quad s.t. \sum_j \gamma_{i,j} = 1 \tag{3}$$

$$\underset{Y}{\text{argmin}} \sum_i \|y_i - \sum_j \gamma_{i,j} y_j\|^2 \tag{4}$$

Here $Y = [y_1, \ldots, y_N] \in R^{3 \times 1 \times t \times N}$. The manifold perspective takes into account the inconsistency between polarimetric scale and time scale. Besides, the intrinsic distribution of data is also explored. Thus, both characteristics of the incoherent feature and coherent feature can be looked after well. Let $C = [c_1, \ldots, c_M] \in R^{3 \times 1 \times t \times M}$ be the dictionary. θ_i is the corresponding code for y_i. Plug y_i into Equation (2) to get the following sparse manifold expression:

$$\underset{\Theta}{\text{argmin}} \sum_{i=1}^N \|y_i - C\theta_i\|^2 + \lambda \|d_i \cdot \theta_i\|^2 \quad s.t. \ 1^T \theta_i = 1, \ \forall i \tag{5}$$

3.2. Compressed Sensing

The construction of the required over-complete dictionary is complex and the sparse feature remains redundant, which yields the performance in some degree. An ideal output code requires low-dimension preservation and high-information retention. For this purpose, we further optimize the sparse priors in Equation (5) with the aim of reducing the signal's dimension by introducing a matrix A in Equation (6).

$$\underset{\Theta}{\text{argmin}} \sum_{i=1}^N \|Ay_i - C\theta_i\|^2 + \lambda \|d_i \cdot \theta_i\|^2 \quad s.t. \ 1^T \theta_i = 1, \ \forall i \tag{6}$$

The compressed sensing [19] method captures and represents compressible signals at a rate significantly below the Nyquist rate. Our presented method takes advantage of compressed sensing for better feature extraction and dimension reduction. The input feature Y is projected into a wavelet base φ. The corresponding coefficients $\alpha = \varphi^T Y$ are found to contain few large values and many small values. So a random Gaussian matrix Φ is introduced to act as an observation matrix for its inconsistency with wavelet basis, i.e., $Z = \Phi\alpha$. This randomness cannot guarantee that the reconstructed feature coding is the sparsest one. Thus, further constraints and optimizations are made in Equation (7).

$$\underset{D,\Theta}{\text{argmin}} \sum_{i=1}^N \|z_i - D\theta_i\|^2 + \lambda \|d_i \cdot \theta_i\|^2 \quad s.t. \ \|d_j\| \leq 1, \ | \ d_{i,j} \geq \xi \ \forall j = 1, \ldots, d, \ i = 1, \ldots, N \tag{7}$$

Here $D = \Phi\varphi^T$. The 2-norm constraint on vector d_i helps to avoid trivial solutions and the constraint on $d_{i,j}$ helps to obtain prominent dictionary atoms.

3.3. Framework

We extract the polarized matrixes from the input images in a same data set and then form their corresponding incoherent feature and coherent feature. After normalization, the features are materials for the construction of the information primitive, which is to be refined by the sparse manifold model. The final feature is trained and tested in a SVM classifier. For contrast experiments of our method, we input the normalized incoherent feature and coherent feature respectively, directly into the same SVM classifier. In addition, the three mentioned fusing methods are also tested. The whole framework of our validation process is shown in Figure 3.
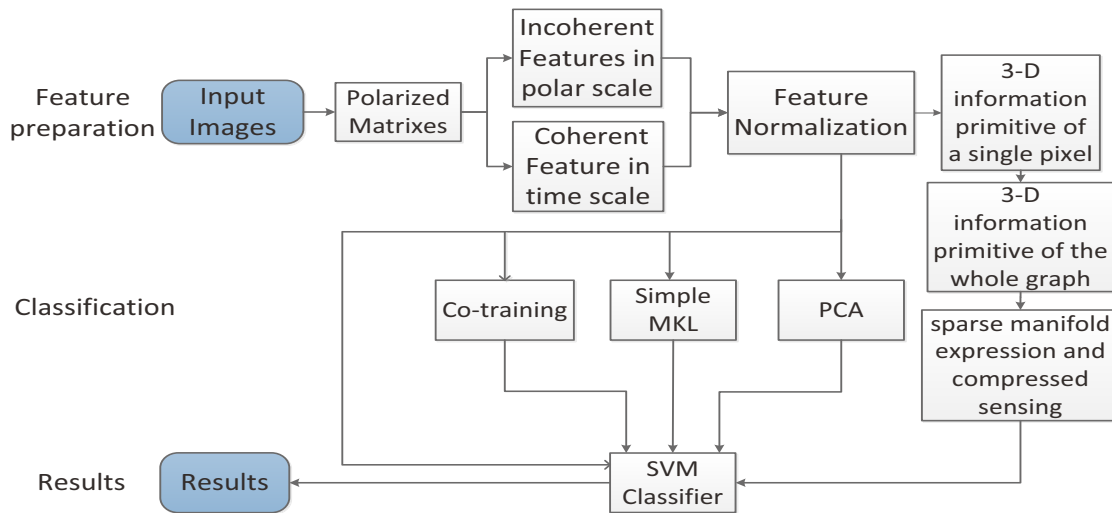
Figure 3. Framework of the validation process.

4. Experiments and Discussion

4.1. Data Sets

The experiments were conducted on three real polarimetric SAR data sets. Details are presented as follows.

4.1.1. Data Set 1

This data set consists of three full polarization SAR images captured by RADARSAT-2 in *Inner Mongolia*, *China* on adjacent dates which are respectively May 23, June 16 and July 19 in 2013. The image size is 5907 × 3572 pixels. Three kinds of land cover categories exist in this area: bare land, forest and farmland.
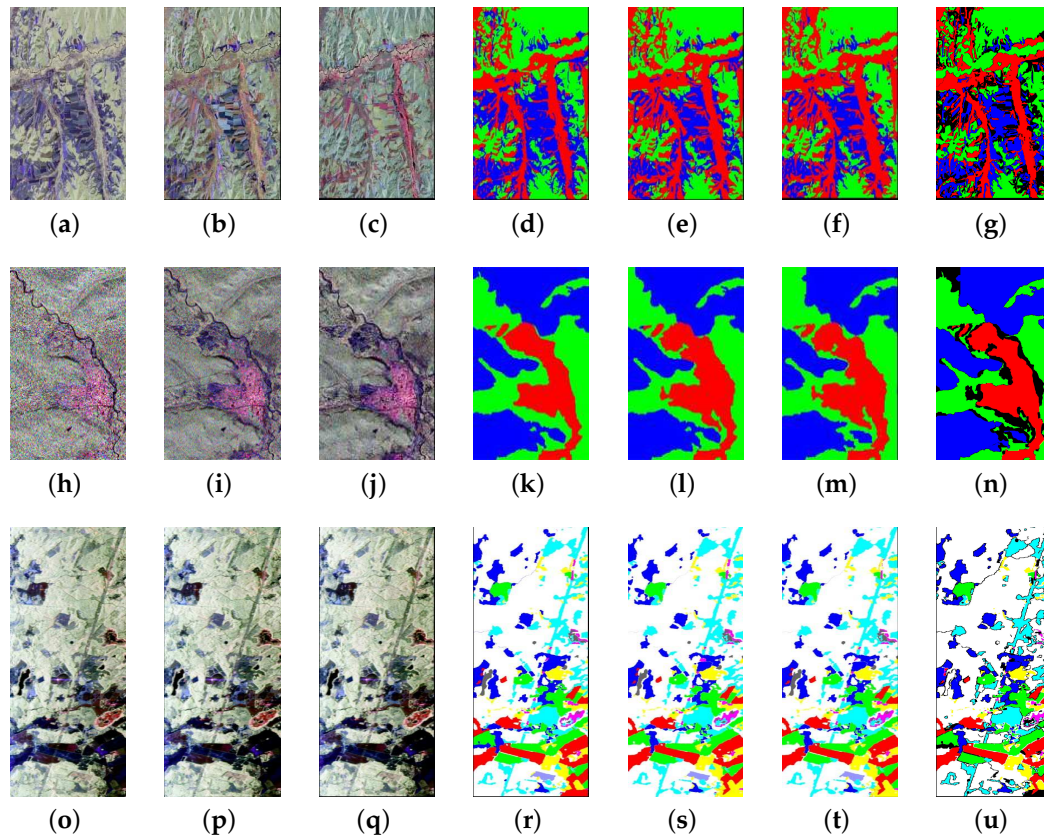
4.1.2. Data Set 2

This data set consists of three full polarization SAR images captured by RADARSAT-2 in *Genhe* city in *Inner Mongolia*, *China* on adjacent dates which are respectively August 20, September 13 and October 7 in 2013. The image size is 1434 × 2050 pixels. Three kinds of land cover categories exist in this area: bare land, forest and building.

4.1.3. Data Set 3

This data set consists of three full polarization SAR images captured by SETHI in *RemingStrop* in *Sweden* on adjacent dates which are respectively August 29, September 9 and September 23 in 2010. The image size is 2372 × 648 pixels. Nine kinds of land cover categories exist in this area: farmland, pine trees, spruces, birches, grassland, bare land, building, water and sand. As the last four kinds of land cover account for only a small proportion, we only take the first five types into consideration.

　　　Based on the ground truths on Google Earth, we marked the label truths by Photoshop according to prior obtained information about how many categories existed in every data set and how these categories were roughly distributed in every testing area. For every mentioned data set, the label truths of the data on three dates were contrasted to extract the shared part, which means that the category of each pixel involved was considered to be constant in the study period. The area not in these parts was labeled in black. The images and their label truths are presented in Figure 4.

Figure 4. (**a**–**c**) present images of data set 1; (**d**–**f**) are their corresponding label truths and (**g**) is the shared label truth. (**h**–**j**) present images of data set 2; (**k**–**m**) are their corresponding label truths and (**n**) is the shared label truth. (**o**–**q**) present images of data set 3; (**r**–**t**) are their corresponding label truths and (**u**) is the shared label truth.

4.2. Experiments and Results Discussion

　　　The pixels in the shared part of each data set were randomly divided into three parts, the first of which accounted for 30% of all the pixels and served as the training set. The second part which consisted of 20% was used as the validation set. The remaining 50% was for testing.

　　　In the stochastic walk period for constructing the coherent feature, the Euclidean measurement between each pair of pixels in the polarimetric decomposition cube was taken as the similarity measurement to calculate the transition probability. The normalization parameter and the scaling parameter were both set as 10, i.e., $K = 10$, $\delta = 10$ for convenience of computation. As for the number of paths m and the number of neighbors k in every path, we used the method in [16] for reference. Here, m and k were required to satisfy $m \geq 2, k \geq 2, 8 < m \times k \leq 26$. The optimal m and k ought to ensure that the paths cover enough neighbors in the 3-D neighborhood and at the same time give consideration to computation cost. We randomly sampled five parameter pairs within the range of m and k and selected the optimal pair which produced the best classification result within the stipulated time. Here, "optimal" was local but our emphasis was not to generate the best coherent feature but to

integrate the incoherent feature with the coherent feature, thus the general performance of stochastic walk was enough for our work. Finally, the selection result was that three paths were arbitrarily taken which separately involved five neighbors in the 3-D neighborhood.

In the sparse manifold model, we trained an over-complete dictionary C with 1024 bases. The constraint parameter λ was set as 0.5. As for the optimal dimension of the observation matrix used in compressed sensing in our model, we conducted a series of analyzing experiments on each validation set. The dimension of the feature generated by sparse manifold expression was 30. However, certain dimensions were found to be contradictory with other dominant dimensions as their existence damaged the overall accuracy. We abandoned these "harmful" dimensions whose damage was beyond an acceptable limit. The remaining X dimensions of the feature were ranked according to their eigenvalues in descending order and then the first x dimensions of the feature were successively taken for analysis. Here, $20 < x \leq \min\{X, 30\}$, as the first 20 components accounted for more than 80 percent of the total information. The optimal choice of each date in each data set can be observed in Figure 5.
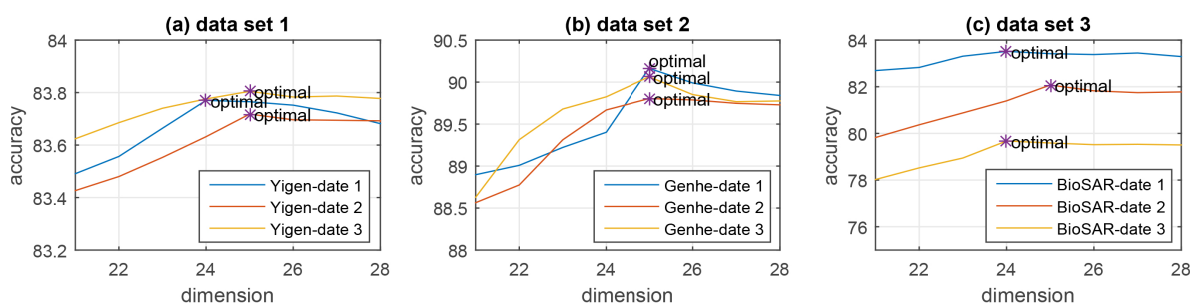
Figure 5. (**a**–**c**) are respectively thetests of the optimal feature dimension on three dates of each data set. Here, "optimal" means that the chosen dimensions of the feature can fully represent the whole feature generated by sparse manifold expression without the sparse feature being redundant.

As a contrast, we also conducted a separate feature classification and the mentioned fusing classifications, namely co-training, Simple MKL and PCA, each followed by a linear SVM to compare with the proposed method.

Table 1 gives the results of experiments on the three data sets. The single number in the "coherent feature" column represents the classification result of the coherent feature constructed by forming the pixels of the first and third date as the 3-D neighborhood of the pixels in the second date. The triplets of numbers in other cells are the average accuracies of the three images of each series. The experiments were conducted on Matlab on a 64 bit Windows 7 system. The corresponding time costs including time for feature preparation, fusing processing and final classification of all the experiments are given in Table 2. Moreover, we reflected back the prediction labels on the first date of each series and recovered the whole graph together with the labels of the training set and validation set. Figure 6 gives the final restored images of each method. The differences between each restored image and the corresponding shared label truth reveal the classification performance of each method.

From the perspective of accuracy, our method has an advantage over the other three feature fusing methods with an average of approximately 7 percent better accuracy. The coherent feature outperforms the incoherent feature with the help of the 3-D neighborhood since the information in time series takes effect other than the information in a single image. Co-training, Simple MKL and PCA obtained better results than the single feature as expected. However, they all combine the coherent and incoherent feature mechanically and the inner connection of features is lost in the structure, thus the fusing results are susceptible to local flaws. Especially in co-training, the feedback process can return the correct information but may also strengthen the validation of the wrong results. The proposed method captures the feature information by a 3-D descriptive primitive in the initial feature construction stage and then optimizes it by a sparse manifold classification model. The primitive integrates the

coherent feature and incoherent feature while the classification model explores the inner distribution of nonlinear SAR data. The information is sufficiently reserved while the dimension is effectively reduced. The fusing results show an obvious advantage.

Table 1. Classification results of three data sets.

Type	Coherent Feature	Incoherent Feature	Fusing Feature			
			Co-training	Simple MKL	PCA	Proposed Method
Accuracy on Data Set 1	73.2777	65.3550	69.4196	73.0301	64.7154	83.7703
		54.0009	63.8058	70.9354	57.9711	83.7181
		48.0356	60.5309	71.1018	61.2217	83.8048
Accuracy on Data Set 2	57.0702	46.5187	67.3346	83.5341	66.2782	90.1641
		46.2600	66.8268	83.7032	66.3161	89.8039
		47.0494	73.4284	83.4196	66.4479	90.0633
Accuracy on Data Set 3	53.8057	44.6433	71.5634	73.6971	52.6820	83.5168
		33.7141	67.1748	73.6069	61.7849	82.0731
		47.0471	73.8655	74.6321	59.0582	79.6735

**Table 2.** Time costs of all the experiments of three data sets.

Type	Coherent Feature	Incoherent Feature	Fusing Feature			
			Co-training	Simple MKL	PCA	Proposed Method
Time Cost of Data Set 1 (m:s)	12:46	12:13	41:23	30:35	28:26	19:17
		13:33	43:07	26:12	27:40	18:13
		12:29	43:18	36:04	33:37	17:34
Time Cost of Data Set 2 (m:s)	11:32	11:15	39:46	26:17	22:14	15:11
		12:01	36:34	25:40	23:31	14:12
		11:41	43:22	24:16	25:05	13:06
Time Cost of Data Set 3 (m:s)	16:07	16:20	53:56	34:09	32:17	22:04
		15:41	58:26	33:50	30:11	23:07
		15:07	53:05	34:21	32:27	21:16

From the perspective of time cost, the proposed method is rather efficient compared with other fusing methods. Single coherent feature classification or incoherent feature classification took less time as their structures were simple and features were low-dimensional but, on the other hand, their performance was less satisfactory. It was worthwhile sacrificing proper time for accuracy in the fusing methods. Co-training took the most time as the disagreement of involved models might only be eliminated by repeated propagations. Simple MKL took a similar amount of time as PCA while both of them were more time-consuming than the proposed method. For Simple MKL, the time was spent on the complex kernel matrix computation while for PCA the time-consuming part was the search of the latent subspace. The time advantage of the proposed method is due to the efficient nonlinear transformation and the sparse feature coding. The sparse manifold expression simplified the feature structure and the compressed sensing processing further lowered the feature dimension. The feature was low-dimensional but discriminatory, thus making the proposed method more efficient.

From the aspect of visual effects, Simple MKL presented the best result in the five contrasted methods. Although it did obtain good results in certain categories, in other categories such as bare land (in red) of the first data set, bare land (in green) of the second data set and spruces (in yellow) of the third data set, it performed poorly. In contrast, our method produced favorable results in all categories involved. The overall performance of the proposed method was superior to that of Simple MKL, which could also be observed from Figure 6.

Moreover, we computed the confusion matrixes of the results achieved by the proposed method on the first date of the three data sets shown in Tables 3 and 4. Observing the diagonals of the matrixes of the first and second data set, the highest accuracy of 96.1337% and 97.7938% was gained for "forest". This is because the appearance of a forest scene is relatively more distinctive than that of other scene themes. The lowest accuracy in the third data set was observed for "birches". The reason is that scenes from the birches in this area are scattered over the grassland and other plants.

Overall, both from the numerical perspective and the visual perspective, the experiments on the three real SAR data sets validate the effectiveness of our presented method. The detailed results of the proposed method also testify its capability to satisfactorily complete the given classification task.
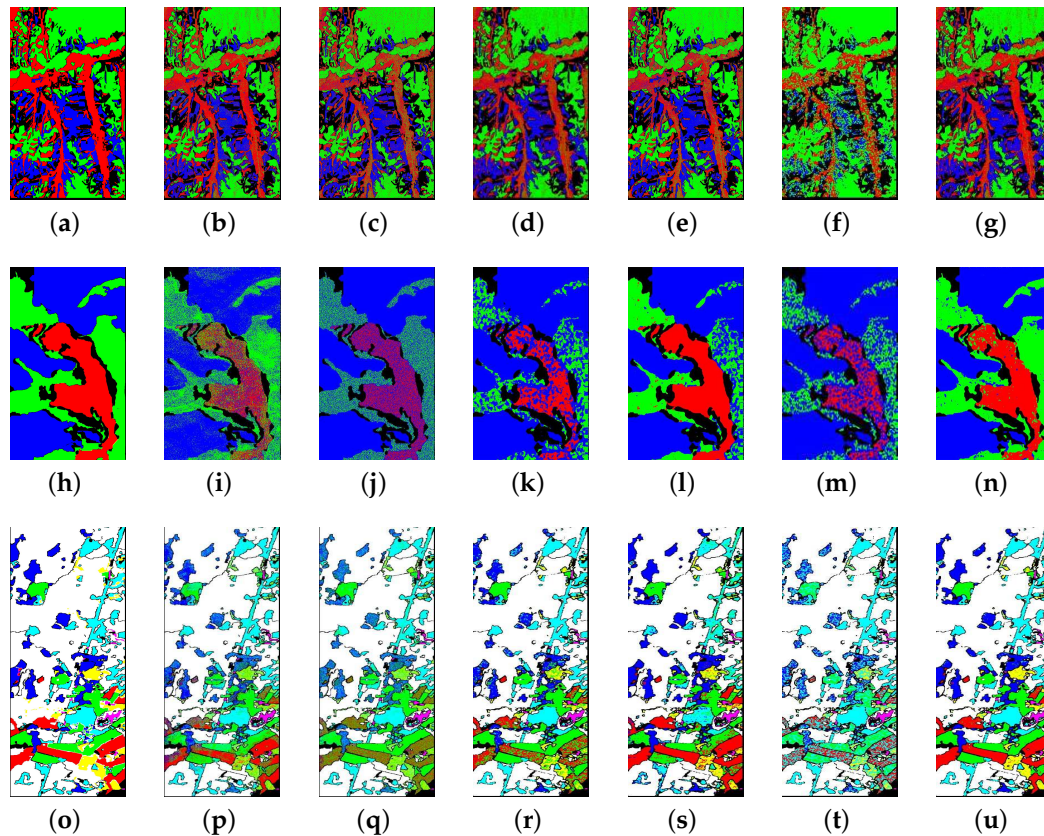
Figure 6. (**a**–**u**) From top to bottom are the shared label truth of each data set and the results on the first date of each data set. From left to right are the shared label truth and the results of coherent feature classification, incoherent feature classification, co-training, Simple MKL, PCA and the proposed method respectively.

Table 3. Confusion matrix of the results achieved by the proposed method on the first date of data set 1 and data set 2.

Data Set 1	Bare Land (Red)	Forest (Green)	Farmland (Blue)
bare land	**91.2493**	4.1328	4.6180
forest	2.6086	**96.1337**	1.2577
farmland	5.2665	3.8476	**90.8860**
Data Set 2	Building (Red)	Bare Land (Green)	Forest (Blue)
building	**87.5113**	10.3495	2.1392
bare land	0.1025	**90.3968**	9.5007
forest	0.0955	2.1107	**97.7938**

Table 4. Confusion matrix of the results achieved by the proposed method on the first date of data set 3.

Data Set 3	Farmland (Red)	Pine Trees (Green)	Spruces (Yellow)	Birches (Cyan)	Grassland (Blue)
farmland	**97.6495**	1.9156	0.2967	0.0153	0.1228
pine trees	0.6353	**98.8160**	0.1959	0.0041	0.3486
spruces	0.2482	0.3911	**97.3556**	0.0134	1.9916
birches	1.5294	6.8386	0.1051	**87.2796**	4.2473
grassland	0.0099	0.1972	0.0657	0.0017	**99.7117**

5. Conclusions

In this paper, we propose a sparse manifold classification method with a multi-dimensional feature on PolSAR image time series.

The proposed method firstly extracts the incoherent feature in the polarimetric scale by target decompositions and extracts the coherent feature in the time scale by stochastic walk, thus constructing a three-dimensional descriptive primitive of a single pixel and further of a whole graph. This approach turns out to be an effective processing foundation. Afterwards, a nonlinear classification model is proposed based on sparse manifold expression and compressed sensing for the purpose of feature extraction and dimension reduction. Finally, the classification is realized with a SVM classifier. The experiment results on three real polarimetric SAR image sets show that our multi-dimensional descriptive primitive can effectively integrate features in the initial feature construction stage which can be a convenient tool for SAR image processing and the nonlinear classification model can also satisfactorily extract information and reduce dimension.

As the proposed method deals with fusion in the early feature construction stage, in future work we intend to proceed to late fusion by combining the results of different classifiers. Under the guidance of the optimal fusion rule of N non-independent detectors [20], late fusion would be cast into the information propagation process for the purpose of identifying the optimal fusion weight for each classifier. For further improvement, we intend to combine the proposed method in this work with late fusion to exert both of their superiorities.

Author Contributions: Chu He and Gong Han conceived and designed the experiments; Di Feng performed the experiments and analyzed the results; Gong Han wrote the paper; Juan Du improved the experiments; Mingsheng Liao revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SAR	Synthetic Aperture Radar
SVM	Support Vector Machine
TD	Target Decomposition
MKL	Multiple Kernel Learning
PCA	Principle Components Analysis
InSAR	Interferometric Synthetic Aperture Radar
LLC	Locality-constrained Linear Coding
LLE	Locally Linear Embedding
PolSAR	Polarimetric Synthetic Aperture Radar

References

1. Jain, A.K.; Duin, R.P.W.; Mao, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 4–37.
2. Richard, H.J. Phenomenological theory of radar targets. Ph.D. Dissertation, Technical University, Delft, The Netherlands, 1970.
3. Chandrasekhar, S. *Radiative Transfer*; Clarendon Press: Oxford, UK, 1950.
4. Robert, C.S.; Eric, P. A review of target decomposition theorems in radar polarimetry. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 498–518.
5. Osmanoglu, B.; Sunar, F.; Wdowinski, S.; Cabral-Cano, E. Time series analysis of InSAR data: Methods and trends. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 90–102.
6. Zhang, Q.; Wu, Y.; Zhao, W.; Wang, F.; Fan, J.; Li, M. Multiple-Scale Salient-Region Detection of SAR Image Based on Gamma Distribution and Local Intensity Variation. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1370–1374.

7.   He, C.; Li, S.; Liao, Z.; Liao, M. Texture Classification of PolSAR Data Based on Sparse Coding of Wavelet Polarization Textons. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4576–4590.

8.   Zhang, L.; Xia, G.S.; Wu, T.; Lin, L.; Tai, X.C. Deep Learning for Remote Sensing Image Understanding. *J. Sens.* **2015**, *501*, doi:10.1155/2016/7954154.

9.   Xu, C.; Tao, D.; Xu, C. A Survey on Multi-view Learning. *Comput. Sci.* **2013**, arXiv:1304.5634v1.

10.   Nigam, K.; Ghani, R. Analyzing the Effectiveness and Applicability of Co-Training. In Proceedings of the Ninth International Conference on Information and Knowledge Management, McLean, VA, USA, 6–11 November 2000; pp. 86–93.

11.   Gonen, M.; Alpaydin, E. Multiple Kernel Learning Algorithms. *J. Mach. Learn. Res.* **2011**, *12*, 2211–2268.

12.   Lu, H.; Plataniotis, K.N.; Venetsanopoulos, A. *Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2013.

13.   Blum, A.; Mitchell, T. Combining Labeled and Unlabeled Data with Co-Training. In Proceedings of the Workshop on Computational Learning Theory, COLT, Madison, WI, USA, 24–26 July 1998; Morgan Kaufmann Publishers: Burlington, MA, USA, 1998; pp. 92–100.

14.   Rakotomamonjy, A.; Bach, F.; Canu, S.; Grandvalet, Y. SimpleMKL. *J. Mach. Learn. Res.* **2008**, *9*, 2491–2521.

15.   Jolliffe, I. *Pincipal Component Analysis*, 2nd ed.; Springer Series in Statistics; Springer: New York, NY, USA, 2002.

16.   Estrada, F.J.; Fleet, D.J.; Jepson, A.D. Stochastic Image Denoising. In Proceedings of the British Machine Vision Conference, BMVC 2009, London, UK, 7–10 September 2009.

17.   Wang, J.; Yang, J.; Lv, F.; Huang, T.; Gong, Y. Locality-Constrained Linear Coding for Image Classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; Volume 119, pp. 3360–3367.

18.   Roweis, S.T.; Saul, L.K. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* **2000**, *290*, 2323–2326.

19.   Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306.

20.   Luis Vergara and Antonio Soriano and Gonzalo Safont and Addisson Salazar. On the fusion of non-independent detectors. *Digit. Signal Process.* **2016**, *50*, 24–33.